

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開2000-10901

(P2000-10901A)

(43)公開日 平成12年1月14日(2000.1.14)

(51)Int.Cl. ⁷	識別記号	F I	テーマコード*(参考)
G 0 6 F 13/14	3 1 0	G 0 6 F 13/14	3 1 0 H 5 B 0 0 5
3/06	5 4 0	3/06	5 4 0 5 B 0 1 4
12/08		12/08	Q 5 B 0 6 5
			B
	3 2 0		3 2 0
審査請求 未請求 請求項の数5 F D (全 12 頁) 最終頁に続く			

(21)出願番号 特願平10-189957

(22)出願日 平成10年6月19日(1998.6.19)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 藤本 和久

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(72)発明者 藤林 昭

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74)代理人 100099302

弁理士 笹岡 茂 (外1名)

最終頁に続く

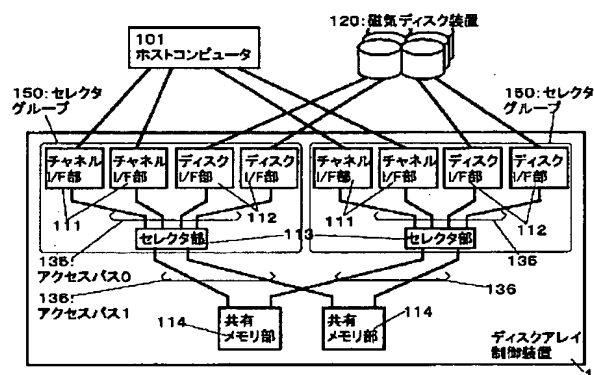
(54)【発明の名称】 ディスクアレイ制御装置

(57)【要約】

【課題】 チャネルI/F部及びディスクI/F部と共有メモリあるいはキャッシュメモリ間のチャネルアクセスパスを効率的に使用し、メモリアクセススループット、特にキャッシュメモリへのアクセススループットを高める。

【解決手段】 チャネルI/F部111、又はディスクI/F部112からは、セレクト部113を介して共有メモリ部114にアクセス可能であり、チャネルI/F部、又はディスクI/F部とセレクト部113間と、セレクト部と共有メモリ部間とは、夫々アクセスパス0 135、アクセスパス1 136により接続され、セレクト部は、チャネルI/F部、又はディスクI/F部からの複数の入力ポートと、共有メモリ部への複数の出力ポートを相互に接続する手段と、複数の入力ポートから出力ポートへの接続要求を接続要求が到着した順に格納する手段と、複数の接続要求間の調停を行い、各出力ポートに出力ポートからの接続要求を割り当てる調停手段から構成される。

図1



【特許請求の範囲】

【請求項 1】 ホストコンピュータとの 1 つ以上のインターフェース部と、複数の磁気ディスク装置との 1 つ以上のインターフェース部と、磁気ディスク装置のデータ及びディスクアレイ制御装置に関する制御情報を格納する物理的に独立した 1 つ以上の共有メモリ部を有し、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部からは、セクタを介して前記共有メモリ部にアクセス可能であり、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部と前記セクタ間と、前記セクタと前記共有メモリ部間はアクセスパスにより接続されているディスクアレイ制御装置であって、前記セクタは、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部からの複数の入力ポートと、前記共有メモリ部への複数の出力ポートを相互に接続する手段と、前記複数の入力ポートから出力ポートへの接続要求を前記接続要求が到着した順に格納する手段と、複数の前記接続要求間の調停を行い、各出力ポートに前記入力ポートからの接続要求を割り当てる調停手段を有しており、前記調停手段は、前記到着順に格納された接続要求の中の先頭の要求が、現在空いている出力ポートへの要求であれば、該要求へ出力ポートを割り当て、前記到着順に格納された接続要求の中の先頭の要求が、現在使用中の出力ポートへの要求であれば、2 番目の要求を調べ、2 番目の接続要求が、現在空いている出力ポートへの要求であれば、該要求へ出力ポートを割り当て、2 番目の接続要求が、現在使用中の出力ポートへの要求であれば、3 番目の要求を調べ、それ以降、多くとも現在空いている出力ポートの数に等しい回数だけ、上記出力ポートへの接続要求の調停（割り当て）を繰り返すことを特徴とするディスクアレイ制御装置。

【請求項 2】 請求項 1 記載のディスクアレイ制御装置において、

前記共有メモリ部は各々物理的に独立した前記共有メモリ部間で二重化されており、前記セクタから前記二重化された共有メモリ部両方への同時アクセスが発生することを特徴とするディスクアレイ制御装置。

【請求項 3】 請求項 1 記載のディスクアレイ制御装置において、

前記共有メモリ部が、磁気ディスク装置のデータを一時的に格納するキャッシュメモリ部と、前記キャッシュメモリ部及び前記ディスクアレイ制御装置に関する制御情報を格納する共有メモリ部に物理的に分割されており、前記キャッシュメモリ部に繋がるセクタと、共有メモリ部に繋がるセクタが物理的に独立しており、

前記ホストコンピュータとのインターフェース部、及び前記複数の磁気ディスク装置とのインターフェース部から、前記キャッシュメモリ部、または前記共有メモリ部へのアクセスパスが物理的に独立しており、

少なくとも前記キャッシュメモリ部に繋がるセクタに前記調停手段を備えることを特徴とするディスクアレイ制御装置。

【請求項 4】 請求項 3 記載のディスクアレイ制御装置において、

10 前記共有メモリ部は各々物理的に独立した前記共有メモリ部間で二重化されており、前記キャッシュメモリ部は各々物理的に独立した前記キャッシュメモリ部間で二重化されており、少なくとも前記キャッシュメモリ部に繋がるセクタから前記二重化されたキャッシュメモリ部両方への同時アクセスが発生し、少なくとも前記キャッシュメモリ部に繋がるセクタに前記調停手段を備えることを特徴とするディスクアレイ制御装置。

【請求項 5】 請求項 1 乃至請求項 4 のいずれかの請求項記載のディスクアレイ制御装置において、

20 前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部から、前記共有メモリ部、または前記キャッシュメモリ部へのアクセスの際に、最初にアドレスとコマンドを連続して送出し、前記共有メモリ部、または前記キャッシュメモリ部までのアクセスパスが確立した後に、データを送出することを特徴とするディスクアレイ制御装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、データを分割して複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置に関する。

【0002】

【従来の技術】 コンピュータの主記憶の I/O 性能に比べて、2 次記憶装置として用いられる磁気ディスク装置を使ったサブシステムの I/O 性能は 3～4 桁程度小さく、従来からこの差を縮めること、すなわちサブシステムの I/O 性能を向上する努力が各所でなされている。サブシステムの I/O 性能を向上させるための 1 つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを分割して複数の磁気ディスク装置に格納する手段、いわゆるディスクアレイと呼ばれるシステムが知られている。

【0003】 例えば、1 つの従来技術（以下、従来技術 1 という）では、図 2 に示すようにホストコンピュータ 101 とディスク制御装置 2 間のデータ転送を実行する複数のチャネル I/F 部 111 と、磁気ディスク装置 120 とディスク制御装置 2 のデータ転送を実行する複数のディスク I/F 部 112 と、磁気ディスク装置 120 のデータを一時的に格納するキャッシュメモリ部 115 と、磁気ディスク 120 のデータ及びディスク制御装

置 2 に関する制御情報を格納する共有メモリ部 114 を備え、キャッシュメモリ部 115 および共有メモリ部 114 は全チャンネル I/F 部 111 及びディスク I/F 部 112 からアクセス可能な構成となっている。この従来技術 1 では、チャンネル I/F 部 111 及びディスク I/F 部 112 と共有メモリ部 114、またチャンネル I/F 部 111 及びディスク I/F 部 112 とキャッシュメモリ部 115 間は 1:1 に接続されている。

【0004】また、他の従来技術（以下、従来技術 2 という）では、図 3 に示すようにホストコンピュータ 101 とディスク制御装置 3 間のデータ転送を実行する複数のチャンネル I/F 部 111 と、磁気ディスク装置 120 とディスク制御装置 3 間のデータ転送を実行する複数のディスク I/F 部 112 と、磁気ディスク装置 120 のデータを一時的に格納するキャッシュメモリ部 115 と、磁気ディスク 120 のデータ及びディスク制御装置 2 に関する制御情報を格納する共有メモリ部 114 を備える。そして、各チャンネル I/F 部 111 及びディスク I/F 部 112 と共有メモリ部 114 間は共有バス 130 で接続され、各チャンネル I/F 部 111 部及びディスク I/F 部 112 とキャッシュメモリ部 115 間は共有バス 130 により接続されている。

【0005】

【発明が解決しようとする課題】ディスクアレイシステムに対する高性能化要求に対して、これまでは、ディスクアレイ制御装置の大規模化や構成要素の高速化、例えばプロセッサ数やキャッシュ容量の増大、高性能プロセッサの適用、内部バス幅の拡大やバス転送能力の向上などで対応してきた。しかしながら、前記従来技術 2 では、内部バスの転送能力がシステムの大規模化および性能向上に追従するのが困難になりつつある。

【0006】そこで内部バス性能を向上させて高いメモリアクセス性能を得るために、前記従来技術 1 のように、プロセッサとメモリ間を 1:1 に接続する方法が考えられる。この方法によれば、メモリに接続したアクセスパス数に比例して内部バス性能が増加する。しかしながら、搭載するプロセッサ数の増大に比例して、共有メモリおよびキャッシュメモリに接続するアクセスパス数も増加する。そのため、内部バス性能を最大限に引き出すために、各プロセッサ-メモリ間のアクセスを効率的に制御する必要がある。

【0007】本発明の目的は上述の課題を解消し、プロセッサ-メモリ間のアクセスパスを効率的に使用し、メモリアクセススループット、特にキャッシュメモリへのアクセススループットの高いディスクアレイ制御装置を提供することにある。

【0008】

【課題を解決するための手段】上記目的を達成するため、本発明は、ホストコンピュータとの 1 つ以上のインターフェース部と、複数の磁気ディスク装置との 1 つ以

上のインターフェース部と、磁気ディスク装置のデータ及びディスクアレイ制御装置に関する制御情報を格納する物理的に独立した 1 つ以上の共有メモリ部を有し、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部からは、セクタを介して前記共有メモリ部にアクセス可能であり、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部と前記セクタ間と、前記セクタと前記共有メモリ部間はアクセスパスにより接続されているディスクアレイ制御装置であって、前記セクタは、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部からの複数の入力ポートと、前記共有メモリ部への複数の出力ポートを相互に接続する手段と、前記複数の入力ポートから出力ポートへの接続要求を前記接続要求が到着した順に格納する手段と、複数の前記接続要求間の調停を行い、各出力ポートに前記入力ポートからの接続要求を割り当てる調停手段を有しており、前記調停手段は、前記到着順に格納された接続要求の中の先頭の要求が、現在空いている出力ポートへの要求であれば、該要求へ出力ポートを割り当て、前記到着順に格納された接続要求の中の先頭の要求が、現在使用中の出力ポートへの要求であれば、2 番目の要求を調べ、2 番目の接続要求が、現在空いている出力ポートへの要求であれば、該要求へ出力ポートを割り当て、2 番目の接続要求が、現在使用中の出力ポートへの要求であれば、3 番目の要求を調べ、それ以降、多くとも現在空いている出力ポートの数に等しい回数だけ、上記出力ポートへの接続要求の調停（割り当て）を繰り返すようにしている。

【0009】また、前記共有メモリ部は各々物理的に独立した前記共有メモリ部間で二重化されており、前記セクタから前記二重化された共有メモリ部両方への同時アクセスが発生するようにしている。

【0010】また、前記共有メモリ部が、磁気ディスク装置のデータを一時的に格納するキャッシュメモリ部と、前記キャッシュメモリ部及び前記ディスクアレイ制御装置に関する制御情報を格納する共有メモリ部に物理的に分割されており、前記キャッシュメモリ部に繋がるセクタと、前記共有メモリ部に繋がるセクタが物理的に独立しており、前記ホストコンピュータとのインターフェース部、及び前記複数の磁気ディスク装置とのインターフェース部から、前記キャッシュメモリ部、または前記共有メモリ部へのアクセスパスが物理的に独立しており、少なくとも前記キャッシュメモリ部に繋がるセクタに前記調停手段を備えるようにしている。

【0011】また、前記共有メモリ部は各々物理的に独立した前記共有メモリ部間で二重化されており、前記キャッシュメモリ部は各々物理的に独立した前記キャッシュメモリ部間で二重化されており、少なくとも前記キャッ

シュメモリに繋がるセクタから前記二重化されたキャッシュメモリ部両方への同時アクセスが発生し、少なくとも前記キャッシュメモリに繋がるセクタに前記調停手段を備えるようにしている。

【0012】また、前記ホストコンピュータとのインターフェース部、または前記複数の磁気ディスク装置とのインターフェース部から、前記共有メモリ部、または前記キャッシュメモリ部へのアクセスの際に、最初にアドレスとコマンドを連続して送出し、前記共有メモリ部、または前記キャッシュメモリ部までのアクセスパスが確

【0013】

【発明の実施形態】本発明の実施例について、以下詳細に説明する。

《実施例1》図1に、本発明の一実施例を示す。ディスクアレイ制御装置1は、チャンネルI/F部111、ディスクI/F部112、セクタ部113、共有メモリ部114と、アクセスパス0135、アクセスパス1136から構成される。

【0014】図13に示すように、チャンネルI/F部111は、ホストコンピュータとの1つのI/F（ホストI/F）51と、1つのマイクロプロセッサ50と、1つの共有メモリアクセス回路52と、共有メモリ部114への1つのアクセスパスI/Fから構成される。データ書き込み時は、ホストI/F51は、ホストコンピュータ101から送られてきたデータをパケットに分割して共有メモリアクセス回路52へ送出する。共有メモリアクセス回路52は、ホストI/F51から送られてくる複数のパケットを1つのアクセスパスを使用して共有メモリ部114へ送出する。データ読み出し時は、共有メモリアクセス回路52は、共有メモリ部114から送られてきた複数のパケットをホストI/F51へ送出する。ホストI/F51は、共有メモリアクセス回路52から送られてきた複数のパケットを1つのデータにまとめ、ホストコンピュータ101へ送出する。マイクロプロセッサ50は、ホストI/F51及び共有メモリアクセス回路52でのデータの送受信を制御する。ディスクI/F部112は、複数の磁気ディスク装置120との1つのI/F（ドライブI/F）、1つのマイクロプロセッサ、共有メモリ部114への1つのアクセス回路と、共有メモリ部114への1つのアクセスパスI/Fから構成されている。図13に示すホストI/F51をドライブI/Fで置き換えた構成となる。データの書き込みおよび読み出し時には、チャンネルI/F部111の説明で述べた処理と少なくとも同様の処理が行われる。ここで、上記に示した個数は一実施例に過ぎず、上記に限定するものではない。

【0015】共有メモリ部114は、磁気ディスク装置120へ記録するデータや、そのデータの管理情報及びシステム情報などの管理情報を格納する。セクタ部1

13には、2つのチャンネルI/F部111、2つのディスクI/F部112からそれぞれ1本ずつ、計4本のアクセスパス0135が接続されている。また、セクタ部113には、2つの共有メモリ部114へのアクセスパス1136が1本ずつ、計2本接続されている。これら1つのセクタ部113とそれに繋がる2つのチャンネルI/F部111及び2つのディスクI/F部112で1つのグループを形成し、セクタグループ150と呼ぶ。本実施例では、ディスクアレイ制御装置1が2つのセクタグループ150を有する。チャンネルI/F部及びディスクI/F部とセクタ部間のアクセスパスと、セクタ部と共有メモリ部間のアクセスパスの間には上記のようなパス数の関係があるため、セクタ部113ではチャンネルI/F部111及びディスクI/F部112からの4本のアクセスパス0135からの要求の内、共有メモリ部114へのアクセスパス1136の数に相当する2個だけを選択して実行する機能を持つ。ここで、上記個数は一実施例に過ぎず、個数を上記に限定するものではない。

【0016】1つのセクタ部113から共有メモリ部114へ接続されるアクセスパスの数を、チャンネルI/F部111及びディスクI/F部112から1つのセクタ部113に接続されるアクセスパスの数より少なくし、チャンネルI/F部111とディスクI/F部112の合計数よりもセクタ部113の数が少なくなるように上記個数を設定すると、共有メモリ部114それぞれへ接続されるアクセスパス数を削減することができる。共有メモリ部のLSIピンネック及びパッケージのコネクタネックという問題が生じた場合、上記のようにすることで、LSIのピンネック及びパッケージのコネクタネックを解消することができる。

【0017】次に、セクタ部113内部の構成について述べる。図4は、セクタ部113内の構成を示している。セクタ部113は、チャンネルI/F部111、またはディスクI/F部112とのI/Fポート210と、共有メモリ部114とのI/Fポート211と、両者間を互いに接続するセクタ206と、I/Fポート210及び211で入出力を行う際のデータのエラーチェック部201と、チャンネルI/F部111、またはディスクI/F部112から送出されたアドレス、コマンド、及びデータをバッファリングするバッファ202と、チャンネルI/F部111、またはディスクI/F部112から送出されたアドレス及びコマンドを解析するアドレス・コマンド（addr、cmd）解析部203と、解析結果を共有メモリ部114とのI/Fポート211への接続要求として、要求の到着順に管理するキュー管理部204と、キュー管理部内に登録された接続要求に基づいてアービトレーションを行い、共有メモリとのI/Fポート211への接続権を決定するアービトレーション部205を有する。

【0018】共有メモリ部のLSIのピンネック及びパッケージのコネクタネックが生じた場合には、前述のように、チャンネルI/F部111、またはディスクI/F部112とのI/Fポート210の数よりも、共有メモリ部114とのI/Fポート211の数を少なくすることで、それらのネックを解消できる。本実施例では、チャンネルI/F部111、またはディスクI/F部112とのI/Fポート210の数を4つ、共有メモリ部114とのI/Fポート211の数を2つとした。

【0019】図12は、アドレス・コマンド（adr、cmd）解析部203、キュー管理部204、及びアービトレーション部205の詳細構成について示している。アドレス・コマンド（adr、cmd）解析部203は、チャンネルI/F部111、またはディスクI/F部112とのI/Fポート210の数に相当する4つのバッファ220を有し、バッファ内に各I/Fポート210からのアドレス（adr）及びコマンド（cmd）を格納する。アドレスは長さが4バイトで最初の1バイトに出力ポート番号（port No.）が示されている。コマンドは長さが4バイトで最初の1バイトにアクセスの種類（読み出し：RD、書き込み：WR、2重読み出し：2R、2重書き込み：2W）が示されている。ここで、共有メモリ部114が2重化されている場合、2重読み出し及び2重書き込みを行う場合がある。このような2重アクセス時には、2つのポートを同時に使用するため、両方のポートの使用権を獲得する必要がある。

【0020】port No. 抽出部221では、アドレスから要求ポート番号を取り出す。本実施例では、port 0に“00”、port 1に“11”を割り当てた。cmd種類抽出部222では、コマンドからアクセスの種類を取り出す。本実施例では、RDに“00”、WRに“01”、2Rに“10”、2Wに“11”を割り当てた。使用ポート決定部223では、アクセスの種類が2重アクセスでない場合port No. をそのまま出力し、2重アクセスの場合それを示す“01”を出力する。キュー管理部204では、アドレス・コマンド（adr、cmd）解析部203から出力されたport No. を到着順に管理テーブル224に登録する。

【0021】アービトレーション部205では、管理テーブル224の先頭から要求port No. を1つ取り出し、バッファ227に格納する。そして、バッファ226に格納された使用中のport No. とバッファ227内の要求port No. を比較器228で比較する。port No. が異なる場合、その番号をセレクト切り替え信号SEL0、SEL1としてセレクト206へ出力し、キュー管理部204内の順番入替え部225へキューの順番を1つ進めるように指示する。port No. が等しい場合、順番入替え部225へキューの順番を入れ替えるように指示する。順番の入れ替

え方法については、図6の調停のフローの説明で述べる。ここで、アドレス、コマンドの長さ、アドレスまたはコマンド内のport No. またはcmd種類がそれぞれ示されている場所、port No. またはcmd種類へのビットの割り当て方は一実施例に過ぎず、上記に限定するものではない。また、共有メモリ部114が2重化されていない場合、2重アクセスは生じないため、cmd種類抽出部222及び使用port決定部223は必要なくなり、port No. 抽出部221の出力をキュー管理部204へ直接入力すればよい。

【0022】次に、セレクト部113での処理の手順について述べる。図5は、チャンネルI/F部111、またはディスクI/F部112とのI/Fポート210のうちの1つにおける処理フローを示している。まず、ステップ301で、チャンネルI/F部111、またはディスクI/F部112内の共有メモリアクセス回路からアクセス要求（REQ ON）が来るまで待機する。アクセス要求が来たら、ステップ302でアドレス（adr）及びコマンド（cmd）を解析する。ステップ303でアドレス（adr）、コマンド（cmd）にエラーがないかどうかチェックし、エラーがあった場合ステップ315でエラー処理を行い、ステップ301のアクセス要求待機状態に戻る。エラーがなかった場合、ステップ304で、共有メモリ部114とのI/Fポート211への接続要求として、キューに登録する。そして、そのキューの内容に基づいてアービトレーションを行う。ステップ305で、要求した共有メモリ部114とのI/Fポート211が獲得できるまで待機する。獲得できたら、ステップ306でセレクト206を切り替えて要求を出したI/Fポート210と獲得したI/Fポート211を接続する。

【0023】次に、ステップ307で共有メモリ（SM）部114へアクセス要求（REQ ON）を出し、アドレス（adr）とコマンド（cmd）を送出する。ステップ308で、共有メモリ部114からアクセス承認（ACK ON）が返ってくるまで待機する。アクセス承認（ACK ON）が返ってきたら、ステップ309でチャンネルI/F部111、またはディスクI/F部112内の共有メモリアクセス回路へアクセス承認（ACK ON）を返す。ステップ310で、データの書き込み時は共有メモリアクセス回路から送られてきたデータを共有メモリ部114へ送信する。また、データの読み出し時は共有メモリ部114から送られてきたデータを共有メモリアクセス回路へ送信する。

【0024】その際、ステップ311でエラーのチェックを行う。エラーを発見した場合、ステップ315でエラー処理を行い、ステップ301のアクセス要求待機状態へ戻る。エラーが無い場合、ステップ312でステータス（Status）が届くのをチェックし、ステータス（Status）が届くまでデータを送信する。ステ

ータス (Status) が届いたら、ステップ 313 で共有メモリ部へ、アクセス承認 (ACK ON) を取り下げるよう指示し、ステップ 301 のアクセス要求待機状態に戻る。

【0025】次に、ステップ 304 でのアービトレーション (調停) の方法について述べる。図 6 は、調停のフローを示している。ステップ 401 で、空いている出力ポートがあるかどうか調べ、空きポートができるまで待つ。ステップ 401 で空きポートがあれば、ステップ 402 でキュー管理部 204 に到着順に格納された接続要求の中の先頭の要求を調べる。ステップ 403 で現在空いている出力ポートへの要求であれば、ステップ 404 でその要求へ出力ポートを割り当てる。ステップ 403 で、キュー管理部 204 に到着順に格納された接続要求の中の先頭の要求が、現在使用中の出力ポートへの要求であれば、ステップ 406 でキュー先頭の要求を (空きポート数+1) 番目にし、ステップ 401 へ戻る。ステップ 404 で出力ポートを割り当てたら、ステップ 405 でキューの順番を 1 つ進め、ステップ 401 へ戻る。上記の制御を行うことにより、共有メモリ部側の I/F

ポート 211 を効率よく割り当てることが可能となり、高いスループットのデータ転送を実現できる。

【0026】また、図 9 に示すように、共有メモリ部 114 を各々物理的に独立した共有メモリ部 114 間で二重化して二重化領域 (160) を形成する。すなわち、2 個の共有メモリ部 114 を 2 重化した場合には各共有メモリ部に同一のデータが書き込まれる。また、各共有メモリ部全体を二重化したり、各共有メモリ部の一部分を二重化したりすることができる。そして、セクタ部 113 から二重化された共有メモリ部 114 両方への同時アクセス (2 重アクセス) が発生するディスクアレイ制御装置 4 では、図 6 のステップ 402、403 においてキュー先頭の要求が 2 重アクセスかどうかを調べ、2 重アクセスの場合、要求される 2 つのポートが空いていればポートを割り当て、そうでなければステップ 406 へ進むという処理を行う。これによって、共有メモリ部 114 に格納したデータの信頼性を向上させることが可能となる。また、磁気ディスク装置 120 に記録するデータの転送時、共有メモリ部 114 との I/F ポート 211 を効率よく割り当てることが可能となり、高いスループットのデータ転送を実現できる。

【0027】《実施例 2》図 1 に示すディスクアレイ制御装置の構成を、図 10 に示すように、共有メモリ部 114 を、磁気ディスク装置 120 に記録するデータを一時的に格納するキャッシュメモリ部 115 と、キャッシュメモリ部 115 及びディスクアレイ制御装置 5 に関する制御情報を格納する共有メモリ部 114 に物理的に分割し、キャッシュメモリ部 115 に繋がるセクタ部 (CMセクタ部) 123 と、共有メモリ部に繋がるセクタ部 (SMセクタ部) 113 を物理的に独立させ

た構成にする。そして、チャンネル I/F 部 111、及びディスク I/F 部 112 から、キャッシュメモリ部 115、または共有メモリ部 114 へのアクセスパス 0135、アクセスパス 1136 を物理的に独立させ、少なくともキャッシュメモリ部 115 に繋がるセクタ部 (CMセクタ部) 123 において、実施例 1 で述べた調停を行う。これは、共有メモリ部に格納されるのは、キャッシュメモリ部 115 及びディスクアレイ制御装置 5 に関する制御情報であり、制御情報のデータ量は少ないので、ポートが使用中となる時間は小さく、すぐにポートは使用可能となるため、調停をしなくても格別の差し障りはないからである。

【0028】また、図 11 に示すように、共有メモリ部 114、及びキャッシュメモリ部 115 を各々物理的に独立した共有メモリ部 114、及びキャッシュメモリ部 115 間でそれぞれ二重化して二重化領域 (160) を形成し、少なくともキャッシュメモリ部に繋がるセクタ部 (CMセクタ部) 123 から二重化されたキャッシュメモリ部 115 両方への同時アクセス (2 重アクセス) が発生するディスクアレイ制御装置 6 では、図 6 のステップ 402、403 においてキュー先頭の要求が 2 重アクセスかどうかを調べ、2 重アクセスの場合、要求される 2 つのポートが空いていればポートを割り当て、そうでなければステップ 406 へ進むという処理を、キャッシュメモリ部に繋がるセクタ部 (CMセクタ部) 123 で行う。これによって、共有メモリ部 114 に格納したデータの信頼性を向上させることが可能となる。また、磁気ディスク装置 120 に記録するデータの転送時、キャッシュメモリ部 115 との I/F ポート 211 を効率よく割り当てることが可能となり、高いスループットのデータ転送を実現できる。

【0029】《実施例 3》図 7 は、チャンネル I/F 部 111 またはディスク I/F 部 112 内の共有メモリ (SM) アクセス回路から共有メモリ部 114 へ、またはチャンネル I/F 部 111 またはディスク I/F 部 112 内のキャッシュメモリ (CM) アクセス回路からキャッシュメモリ部 115 へデータの書き込みを行うときの処理の流れを示している。データの書き込み時は、ステップ 501 で SM または CM アクセス回路からセクタ部 113 または 123 へアクセス要求 (REQ) を出し、続けてステップ 502、503 でアドレス (ADR)、コマンド (CMD) を送出する。ステップ 504、505 でセクタ部 113 または 123 において、アービトレーションを行い、セクタを切り替えて共有メモリ部 114 またはキャッシュメモリ部 115 へのポートを割り当てる。ステップ 506 で、セクタ部 113 または 123 から共有メモリ部またはキャッシュメモリ部へアクセス要求 (REQ) を出し、続けてステップ 507、508 でアドレス (ADR)、コマンド (CMD) を送出する。

【0030】ステップ509で、共有メモリ部114またはキャッシュメモリ部115において、アクセスするメモリモジュールの選択を行い、選択後、ステップ510でセクタ部113または123を経由してSMまたはCMアクセス回路へアクセス承認(ACK ON)を返す。SMまたはCMアクセス回路はACK ONを受けたら、ステップ511でデータを送出する。共有メモリ部114またはキャッシュメモリ部115はデータを全て受け取ったら、ステップ512で後処理を行い、ステップ513でセクタ部113または123を経由してSMまたはCMアクセス回路へステータス(STATUS)を返す。セクタ部113または123はSTATUSを受けたら、ステップ514で共有メモリ部114またはキャッシュメモリ部115へ、アクセス承認を取り下げるように指示を出す(ACK OFF)。SMまたはCMアクセス回路はSTATUSを受けたら、ステップ515でセクタ部113または123に、アクセス承認を取り下げるように指示を出す(ACK OFF)。

【0031】図8は、チャンネルI/F部111またはディスクI/F部112内の共有メモリアクセス回路へ共有メモリ部114から、またはチャンネルI/F部111またはディスクI/F部112内のキャッシュメモリアクセス回路へキャッシュメモリ部115からデータの読み出しを行うときの処理の流れを示している。データの読み出し時の処理ステップ601～610までは、データの書き込み時の処理ステップ501～510までと同じである。

【0032】その後、共有メモリ部114またはキャッシュメモリ部115において、ステップ611で読み出しの前処理を行う。ステップ612で、セクタ部113または123を経由してSMまたはCMアクセス回路へデータを送出する。データの送出自ら終了したら、共有メモリ部114またはキャッシュメモリ部115において、ステップ613で後処理を行い、ステップ614でセクタ部113または123を経由してSMまたはCMアクセス回路へSTATUSを返す。セクタ部113または123はSTATUSを受けたら、ステップ615で共有メモリ部114またはキャッシュメモリ部115へ、アクセス承認を取り下げるように指示を出す(ACK OFF)。SMまたはCMアクセス回路はSTATUSを受けたら、ステップ616でセクタ部113または123へ、アクセス承認を取り下げるように指示を出す(ACK OFF)。

【0033】上記のように、チャンネルI/F部111、またはディスクI/F部112から、共有メモリ部114またはキャッシュメモリ部115へのアクセスの際に、最初にアドレスとコマンドを連続して送出し、共有メモリ部114、またはキャッシュメモリ部115までのアクセスパスが確立した(ステップ510または61

0) 後データを送出することによって、セクタ部113または123において、転送データをバッファリングする必要がなくなり、セクタ部113または123での制御が簡単化され、メモリへのアクセススループットの向上が可能となる。

【0034】

【発明の効果】本発明によれば、ホストコンピュータとのインターフェース部、または複数の磁気ディスク装置とのインタフェース部と共有メモリ部の間にあるセクタ部において、ホストコンピュータとのインターフェース部、または複数の磁気ディスク装置とのインタフェース部から共有メモリ部へのアクセス要求を、共有メモリ部へのアクセスパスへ効率よく振り分けることが可能となる。それによって、ディスクアレイ制御装置のデータ転送のスループットを向上することができる。

【図面の簡単な説明】

【図1】本発明によるディスクアレイ制御装置の構成を示す図である。

【図2】従来のディスクアレイ制御装置の構成を示す図である。

【図3】従来のディスクアレイ制御装置の構成を示す図である。

【図4】本発明によるディスクアレイ制御装置内のセクタ部の構成を示す図である。

【図5】セクタ部での動作フローを示す図である。

【図6】セクタ部内のアービトレーション部での動作フローを示す図である。

【図7】共有メモリ部またはキャッシュメモリ部へデータの書き込みを行うときのシーケンスを示す図である。

【図8】共有メモリ部またはキャッシュメモリ部からデータの読み出しを行うときのシーケンスを示す図である。

【図9】本発明によるディスクアレイ制御装置の他の構成を示す図である。

【図10】本発明によるディスクアレイ制御装置の他の構成を示す図である。

【図11】本発明によるディスクアレイ制御装置の他の構成を示す図である。

【図12】本発明によるディスクアレイ制御装置内のセクタ部の部分の詳細構成を示す図である。

【図13】チャンネルI/F部の構成を示す図である。

【符号の説明】

1、4、5、6 ディスクアレイ制御装置

50 マイクロプロセッサ

51 ホストI/F

52 共有メモリアクセス回路

101 ホストコンピュータ

111 チャンネルI/F部

112 ディスクI/F部

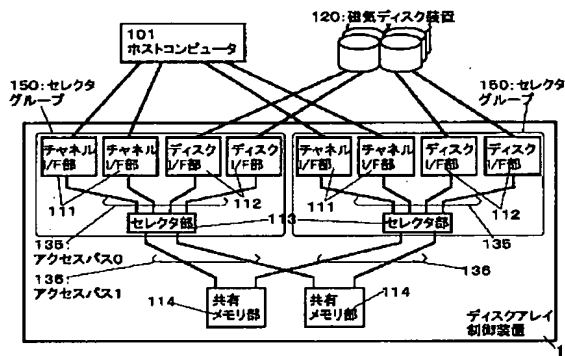
113 セクタ部(SMセクタ部)

13

- 114 共有メモリ部
- 115 キャッシュメモリ部
- 120 磁気ディスク装置
- 123 CMセクタ部

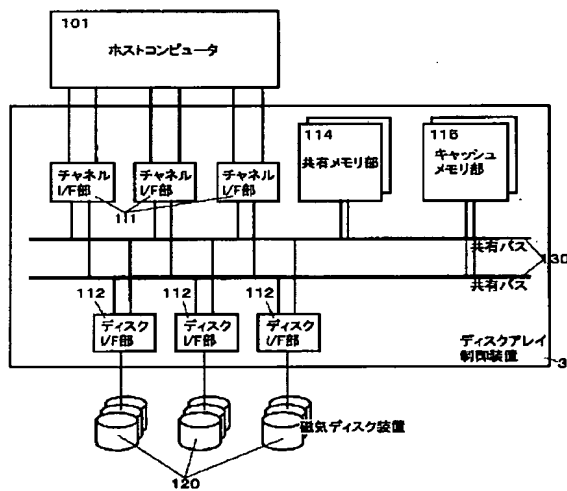
【図1】

図1



【図3】

図3

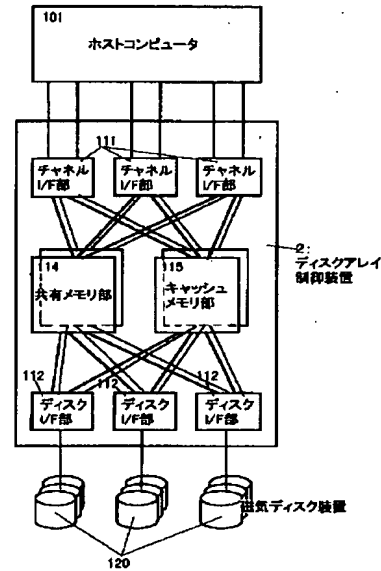


14

- 135 アクセスパス0
- 136 アクセスパス1
- 150 セクタグループ

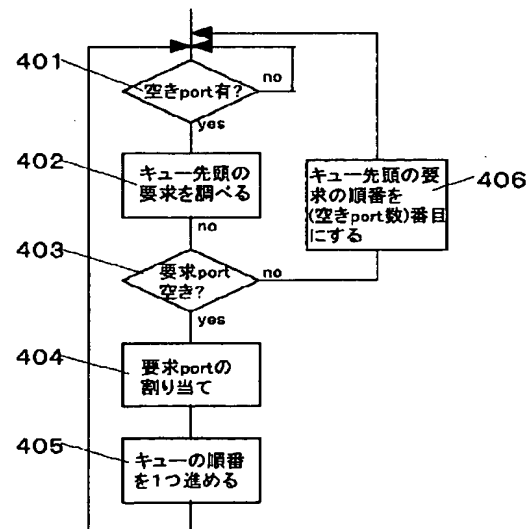
【図2】

図2



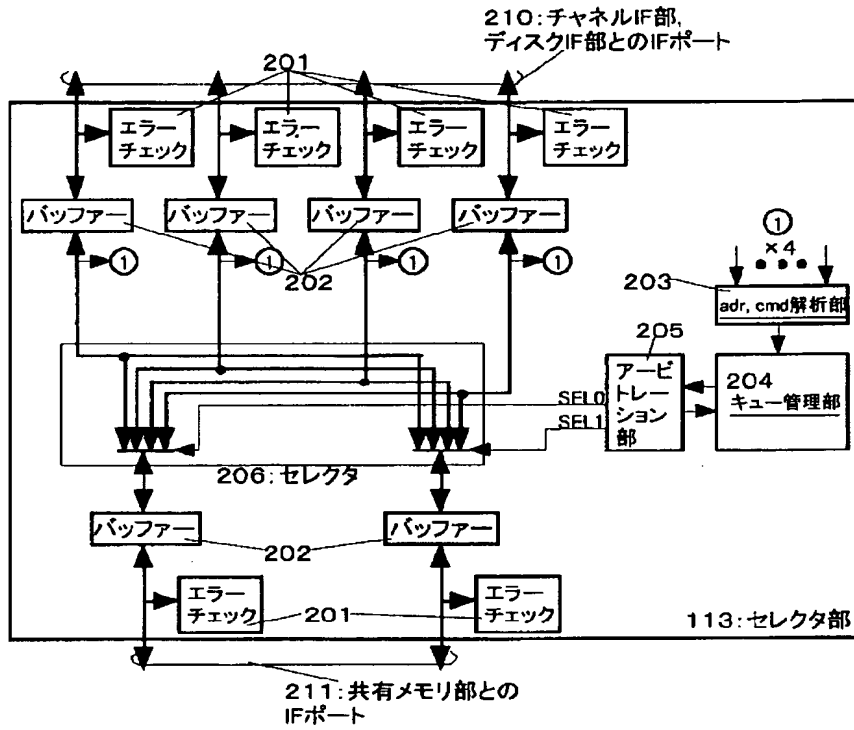
【図6】

図6



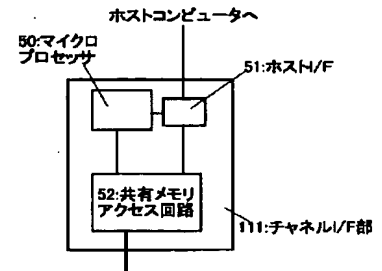
【図 4】

図 4



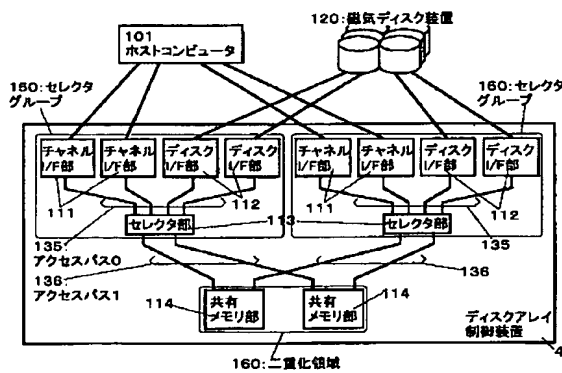
【図 13】

図 13



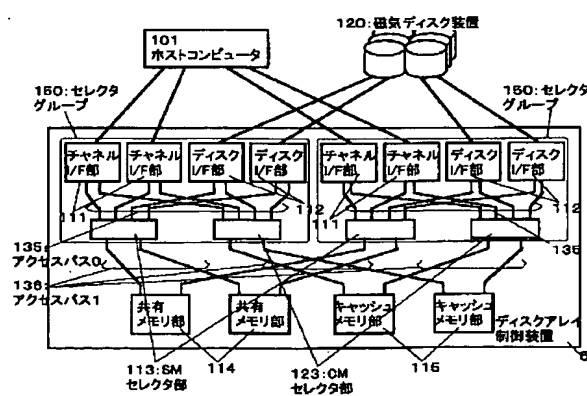
【図 9】

図 9



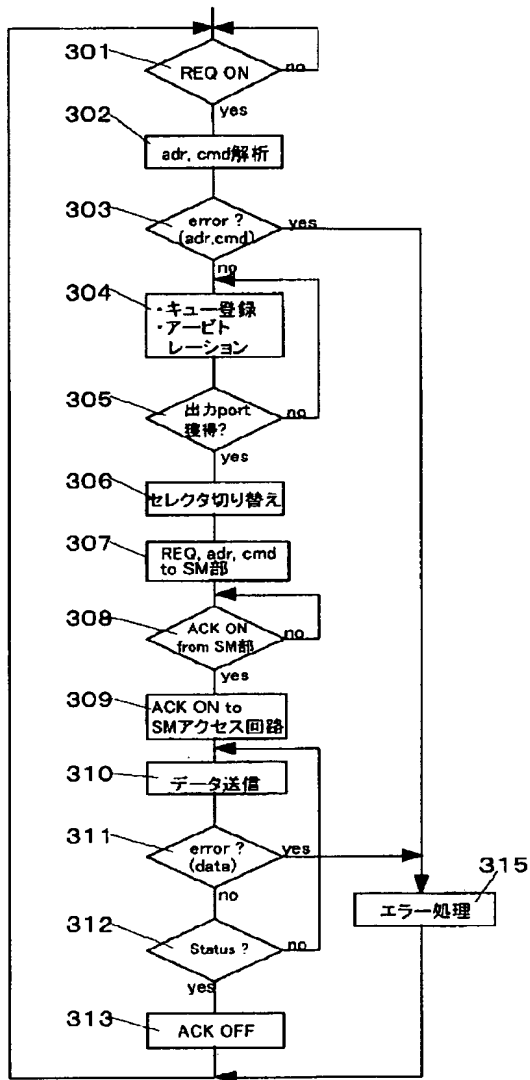
【図 10】

図 10



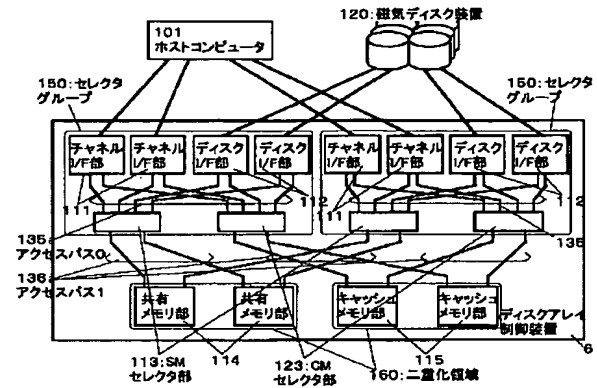
【図 5】

図 5



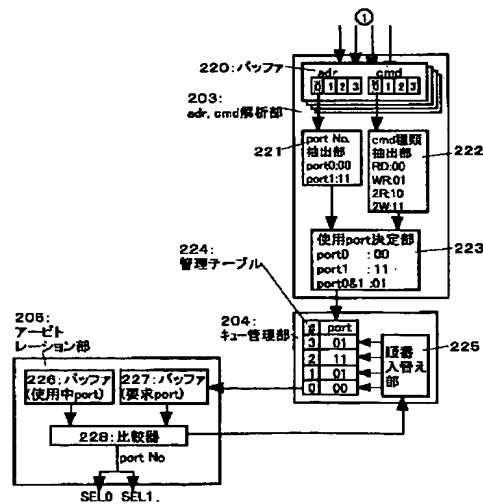
【図 11】

図 11



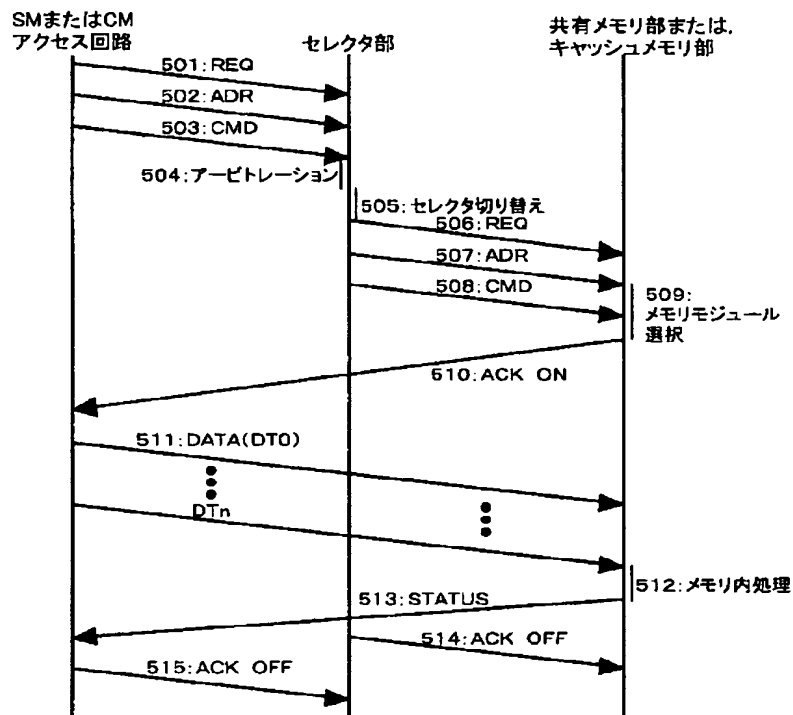
【図 12】

図 12



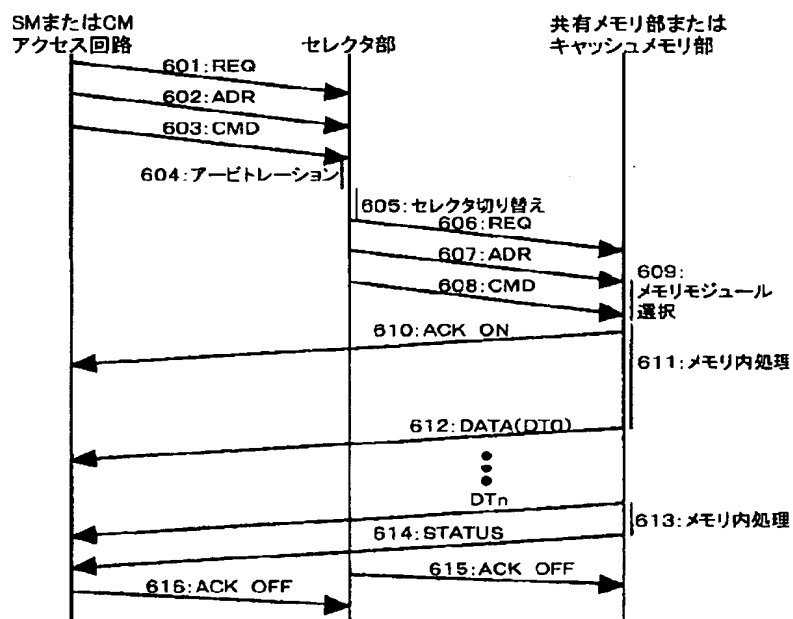
【図 7】

図 7



【図 8】

図 8



フロントページの続き

(51) Int. Cl.⁷

識別記号

F I

テーマコード (参考)

G 0 6 F 13/12

3 3 0

G 0 6 F 13/12

3 3 0 T

F ターム (参考) 5B005 JJ11 KK12 KK15 MM11 NN12

NN71

5B014 EB05 GC36 HA11

5B065 BA01 CA03 CA30 CE11